RFID Cardinality Estimation with Blocker Tags

Xiulong Liu*†, Bin Xiao†, Keqiu Li*, Jie Wu‡, Alex X. Liu¶§, Heng Qi* and Xin Xie*
*School of Computer Science and Technology, Dalian University of Technology, China
†Department of Computing, The Hong Kong Polytechnic University, Hong Kong
‡Department of Computer and Information Sciences, Temple University, USA
¶Department of Computer Science and Engineering, Michigan State University, USA
§National Key Laboratory for Novel Software Technology, Nanjing University, China
Email: {xiulongliudut, likeqiu, qhclement, xiexin0211}@gmail.com
csbxiao@comp.polyu.edu.hk, jiewu@temple.edu, alexliu@nju.edu.cn

***Abstract*—The widely used RFID tags impose serious privacy concerns as a tag responds to queries from readers no matter they are authorized or not. The common solution is to use a commercially available blocker tag which behaves as if a set of tags with known blocking IDs are present. The use of blocker tags makes RFID estimation much more challenging as some genuine tag IDs are covered by the blocker tag and some are not. In this paper, we propose REB, the first RFID estimation scheme with the presence of blocker tags. REB uses the framed slotted Aloha protocol specified in the C1G2 standard. For each round of the Aloha protocol, REB first executes the protocol on the genuine tags and the blocker tag, and then virtually executes the protocol on the known blocking IDs using the same Aloha protocol parameters. The basic idea of REB is to conduct statistically inference from the two sets of responses and estimate the number of genuine tags. We conduct extensive simulations to evaluate the performance of REB, in terms of time-efficiency and estimation reliability. The experimental results reveal that our REB scheme runs tens of times faster than the fastest identification protocol with the same accuracy requirement.**

***Keywords*—RFID Estimation, RFID Privacy, Blocker Tags.**

I. Introduction

RFID systems have been widely used in a variety of applications such as supply chain management and inventory control [1]–[7] as the cost of commercial passive RFID tags is negligible compared with the value of the products to which they are attached (*e.g.*, as low as 5 cents per tag [8]). For example, in Hong Kong International Airport where RFID systems are used to track shipment, the average daily cargo tonnage in May 2010 was 12K tonnes and has been on the rise [9]. An RFID system typically consists of a reader and a population of tags [10]. A reader has a dedicated power source with significant computing capability. It transmits commands to query a set of tags and the tags respond over a shared wireless medium. A tag is a microchip with an antenna in a compact package that has limited computing capability and longer communication range than barcodes. There are two types of tags: (1) passive tags, which do not have their own power sources and are powered up by harvesting the radio frequency energy from readers; (2) active tags, which have their own power sources.

The widely used RFID tags impose serious privacy concerns as when a tag is interrogated by an RFID reader, no matter the reader is authorized or not, it blindly responds with its ID and other stored information (such as manufacturer, product type, and price) in a broadcast fashion. For example, a woman may not want her dress sizes and a patient may not want his/her medication, to be publicly known. An effective solution to this privacy issue is to use commercially available *blocker tags* [11], [12]. A *blocker tag* is an RFID device that is preconfigured with a set of known RFID tag IDs, which we call *blocking IDs*. The blocker tag behaves as if all tags with its blocking IDs are present. A blocker tag protects the privacy of the set of genuine tags whose IDs are among the blocking IDs of the blocker tag because any response from a genuine tag is coupled with the simultaneous response from the blocker tag; thus, the two responses collide and attackers cannot obtain private information.

This paper concerns with the problem of RFID (population size) estimation with the presence of a blocker tag. Formally, the problem is defined as follows: *given (1) a set of unknown genuine tags G of unknown size g, (2) a blocker tag with a set of known blocking IDs B, (3) a required confidence interval $\alpha \in (0, 1]$, and (4) a required reliability $\beta \in [0, 1)$, we want to use one or more readers to compute the estimated the number of genuine tags in G, denoted as $\hat{g}$, so that $P\{|\hat{g} - g| \leq \alpha g\} \geq \beta$.* In other words, we have a set G of genuine tags with unknown number of unknown IDs and a set B of tags with known number of known IDs, we want to estimate $|G|$ with the presence of B. The two sets G and B may overlap, as shown in Fig. 1. This problem may arise in many applications. For example, a jewel store may want to use such an RFID estimation scheme to monitor its stock while a blocker tag is being used to protect the privacy of some precious items.

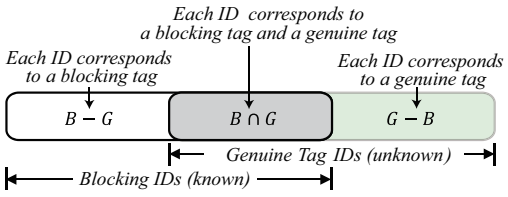

Fig. 1. Three types of IDs in the system containing blocker tags.

To the best of our knowledge, this paper is the first to investigate RFID estimation with the presence of a blocker tag. Although some RFID estimation schemes have been proposed [7], [10], [13]–[18], none of them considers the

978-1-4799-8381-0/15/$31.00 ©2015 IEEE

presence of a blocker tag. Furthermore, none of them can be easily adapted to solve our problem. How about turning off the blocker tag and then using prior RFID estimation schemes to estimate the number of genuine tags? Turning off the blocker tag will give attackers a time window to breach privacy, especially for the scenarios that RFID estimation schemes are being continuously performed for monitoring purpose. Existing tree walking based [19] and framed slotted Aloha based [20] RFID identification schemes can be used to exactly identify the genuine tags, and thus obtaining the genuine tag cardinality. However, they are too slow for our estimation purpose.

In this paper, we propose an *RFID Estimation scheme with Blocker tags* (*REB*). The communication protocol used by REB is the standard framed slotted Aloha protocol, in which a reader first broadcasts a value f and a random number R to the tags where f represents the number of time slots in a forthcoming frame. Then, each tag computes a hash using the random number R and its ID, where the resulting hash value h is within $[0, f-1]$, and the tag replies during slot h. For each slot, if no tag replies, we represent it as 0; if only one tag replies, we represent it as 1; if more than one tag replies, the tag responses will collide, and we represent this slot as c. Note that a reader can detect if there is a collision according to the C1G2 standard. Executing this protocol for the blocking IDs (simulated by the blocker tag) and genuine tags, we get a ternary array $\mathbb{BG}[0..f-1]$ where each bit is 0, 1, or c. As we know the blocking IDs, we can virtually execute the framed slotted Aloha protocol using the same frame size f and random number R for the blocking IDs; thus, we get a ternary array $\mathbb{B}[0..f-1]$ where each bit is 0, 1, or c. From the two arrays $\mathbb{BG}[0..f-1]$ and $\mathbb{B}[0..f-1]$, we calculate two numbers: N_{00}, which is the number of slots i such that both $\mathbb{BG}[i]=0$ and $\mathbb{B}[i]=0$, and N_{11}, which is the number of slots i such that both $\mathbb{BG}[i]=1$ and $\mathbb{B}[i]=1$. REB is based on the key insight that in general the smaller N_{00} is, the larger $|B \cup G|$ is and the larger N_{11} is, the larger $|B-G|$ is. In this paper, we show that N_{00} monotonously decreases with the increase of $|B \cup G|$ and N_{11} monotonously increases with the increase of $|B-G|$. Thus, from the observed N_{00} and N_{11}, we can estimate $|B \cup G|$ and $|B-G|$. Then, we can calculate the size of G because $|G| = |B \cup G| - |B-G|$.

We make the following three key contributions in this paper. First, we make the first effort towards RFID estimation with the presence of a blocker tag. We propose the REB scheme jointly using N_{00} and N_{11} to achieve an unbiased estimator for the genuine tag cardinality. The key technical development of this paper is on quantitatively and statistically correlating N_{00} and $|B \cup G|$, N_{11} and $|B-G|$. Second, we conduct thorough analysis to optimize system parameters, thereby achieving the required confidence interval and reliability in the fastest speed. Third, we implement REB in Matlab and evaluate its performance through extensive simulations. The experimental results reveal that our REB scheme runs tens of times faster than the fastest identification protocol under the same accuracy requirement.

The rest of this paper is organized as follows. In Section II, we describe REB and our theoretical analysis. In Section III, we conduct extensive simulations to evaluate the performance of REB. We discuss related work in Section IV. Finally, we conclude the paper in Section V.

II. REB Protocol

In this section, we first describe the system model used in this paper. Then, an efficient *RFID Estimation scheme with Blocker tags (REB)* is proposed to estimate the number of genuine tags by jointly using N_{00} and N_{11} observed in a time frame. We explicitly give the functional estimator and point out that the estimation using a single time frame is hard to be accurate due to probabilistic variance. Hence, we propose to use multiple independent time frames to refine the estimation. This section further presents rigorous theoretical analysis to investigate how many frames are needed to guarantee the desired estimation accuracy and how to avoid premature protocol termination. We also investigate the parameter settings (*i.e.*, f and p) to optimize the performance of our REB.

A. System Model

For the clarity of presentation, we first consider the RFID system containing a single reader, a single blocker tag, and a population of genuine tags. Then, we will discuss how to extend REB to the scenario that deploys multiple readers and blocker tags. We represent the set of blocking IDs as B, whose cardinality is b. The set of genuine tags is denoted as G, whose cardinality is g. We use U to denote the union tag set, *i.e.*, $B \cup G$, and $|U| = u$. The IDs in $B-G$ do not correspond to any genuine tags, whose cardinality is denoted as b', *i.e.*, $b' = |B-G|$.

The reader communicates with tags (including both genuine tags and virtual ones *simulated* by the blocker tag) under control of the backend server. The communication between the reader and tags are based on a time slotted way. Any two consecutive transmissions (from a tag to a reader or vice versa) are separated by a waiting time $\tau_w = 302us$ [10]. According to the specification of the Philips I-Code system [21], the wireless transmission rate from a tag to a reader is $53Kb/s$, that is, it takes a tag $\tau_t = 18.9us$ to transmit 1 bit. The rate from a reader to a tag is $26.5Kb/s$, that is, transmission of 1 bit to tags requires $\tau_r = 37.7us$. Then, the time of a slot for transmitting m-bit information from a tag to the reader is $\tau_w + m \times \tau_t$; and the time of a slot for transmitting m-bit information from a reader to the tags is $\tau_w + m \times \tau_r$. The notations used throughout the paper are summarized in Table I.

B. Protocol Description

Our REB uses the standard framed slotted Aloha protocol specified in EPC C1G2 [22] as the MAC layer communication mechanism. The reader initializes a slotted time frame by broadcasting a binary request $\langle R, f \rangle$, where R is a random number and f is the frame size (*i.e.*, the number of slots in the forthcoming frame). Using the received parameters $\langle R, f \rangle$, each tag initializes its slot counter sc by calculating $sc = H(ID, R) \mod f$ and the hashing result follows a uniform distribution within $[0, f-1]$. The reader broadcasts *QueryRep* command at the end of each slot. Upon receiving *QueryRep*, a tag decrements its slot counter sc

TABLE I. NOTATIONS USED IN THE PAPER

Notations	Descriptions
G	set of genuine tags.
g	cardinality of G. *i.e.*, $g = \|G\|$.
B	set of blocking IDs.
b'	cardinality of $B - G$. *i.e.*, $b' = \|B - G\|$.
U	union set. $U = B \cup G$.
u	cardinality of U. *i.e.*, $u = \|B \cup G\|$.
α	required confidence interval.
β	required reliability.
$\hat{g}$	estimate of g.
f	frame size.
p	persistence probability.
$E(\cdot)$	expectation.
$Var(\cdot)$	variance.
Z_β	the percentile of β. *e.g.*, $Z_\beta = 1.96$ when $\beta = 95\%$.
p_{00}	probability that a slot pair is $\langle 0, 0\rangle$.
p_{11}	probability that a slot pair is $\langle 1, 1\rangle$.
N_{00}	# of the persistent empty slots in a frame.
N_{11}	# of the persistent singleton slots in a frame.

by 1. In a slot, a tag will respond to the reader if its slot counter sc becomes 0. According to the occupation status, slots are classified into three types: *empty slot* in which no tag responds; *singleton slot* in which only one tag responds; *collision slot* in which two or more tags respond.

In the following, we present how our REB estimates the number of genuine tags by observing the slots in a frame. Since the backend server gets full knowledge of the simulated blocking IDs, it is able to predict which slots the blocking IDs are "mapped" to. Thus, it is able to construct a virtual ternary array $\mathbb{B}[0..f-1]$. A bit in $\mathbb{B}[0..f-1]$ is set to 0 when no blocking ID is mapped to this slot; 1 when only one blocking ID is mapped to this slot; c when two or more blocking IDs are mapped to this slot (a hashing collision). On the other hand, by observing the frame, the reader could get another array $\mathbb{BG}[0..f-1]$, also consisting of f bits. A bit in $\mathbb{BG}[0..f-1]$ is set to 0 when no tag responds in this slot; 1 when only one tag responds in this slot; c when two or more tags cause a collision in this slot. To distinguish a singleton slot from a collision one, each tag does not need to respond with the whole 96-bit ID. For saving time, each tag responds with the RN16 (16-bit) [22] that is much shorter than 96-bit ID. Two slots with the same index in $\mathbb{B}[0..f-1]$ and $\mathbb{BG}[0..f-1]$ are called a slot pair. In our scheme, the reader needs to record the numbers of the following two types of slot pairs.

- N_{00} is the number of *persistent-empty* slot pairs $\langle 0, 0\rangle$ (*i.e.*, $\mathbb{B}[i] = 0$ *AND* $\mathbb{BG}[i] = 0$, $i \in [0, f-1]$).
- N_{11} is the number of *persistent-singleton* slot pairs $\langle 1, 1\rangle$ (*i.e.*, $\mathbb{B}[i] = 1$ *AND* $\mathbb{BG}[i] = 1$, $i \in [0, f-1]$).

REB can estimate the cardinality of genuine tags by jointly using the number of persistent-empty slots and that of persistent-singleton slots. A persistent-empty slot happens only when no ID in $U = B \cup G$ is mapped to this index. Thus, N_{00} reflects the cardinality u of U. Latter, we will show that a *monotone* functional relationship can be established between u and N_{00}. REB uses this function to estimate u from N_{00}. Similarly, a persistent-singleton slot happens when only one ID in $B-G$ is mapped to this index. Therefore, N_{11} reflects the cardinality $|B - G|$ (denoted as b'). Clearly, if we know u and b', we can get the cardinality g of genuine tags by calculating $g = u - b'$. It may not be sufficient to satisfy the required estimate accuracy by counting the numbers of N_{00} and N_{11} in a *single* frame. To improve the accuracy, REB requires the reader to execute k independent frames with different random number R.

Note that, the frame size should be set no more than 512 in practice [10], [19], [23] (the detailed reasons can be found in literature [19]). If a large number of tags contend for such a short frame, most slots will become collision slots. To scale to a large tag population, we exploit the method stated in [10]. Specifically, the reader uses a persistence probability $p \in (0, 1]$ to *virtually* extends the frame size f to f/p, but *actually* terminates the frame after the first f slots. Fundamentally, each tag participates in the actual frame of f slots with a probability p.

C. Functional Estimator

In this section, we derive the functional estimator $\hat{g}$ from N_{00} and N_{11} for the REB protocol in one frame. For an arbitrary slot pair, the probability that it is $\langle 0, 0\rangle$, denoted as p_{00}, is given as follows.

$$p_{00} = (1 - \frac{p}{f})^u \approx e^{-\frac{up}{f}} \tag{1}$$

The approximation in Eq. (1) holds when f/p is relatively large [5], [10], [13]. The number of slot pairs $\langle 0, 0\rangle$, *i.e.*, N_{00}, follows $Bernoulli(f, p_{00})$. The expectation and variance of the variable N_{00} are presented as follows.

$$E(N_{00}) = f \times p_{00} = fe^{-\frac{up}{f}} \tag{2}$$

$$Var(N_{00}) = f \times p_{00} \times (1 - p_{00}) = fe^{-\frac{up}{f}}(1 - e^{-\frac{up}{f}}) \tag{3}$$

Similarly, we use p_{11} to denote the probability that a slot pair is $\langle 1, 1\rangle$, which is given as follows.

$$p_{11} = \binom{b'}{1}(\frac{p}{f})(1 - \frac{p}{f})^{u-1} \approx \frac{b'p}{f}e^{-\frac{up}{f}} \tag{4}$$

The number of $\langle 1, 1\rangle$ slot pairs, *i.e.*, N_{11}, also follows $Bernoulli(f, p_{11})$. The expectation and variance of the variable N_{11} are presented as follows.

$$E(N_{11}) = f \times p_{11} = b'pe^{-\frac{up}{f}} \tag{5}$$

$$Var(N_{11}) = f \times p_{11} \times (1 - p_{11}) = b'pe^{-\frac{up}{f}}(1 - \frac{b'p}{f}e^{-\frac{up}{f}}) \tag{6}$$

According to Eq. (2), u can be expressed as follows.

$$u = -\frac{f}{p}\ln[\frac{E(N_{00})}{f}] \tag{7}$$

Dividing Eq. (5) by Eq. (2), we have:

$$\frac{E(N_{11})}{E(N_{00})} = \frac{b'p}{f} \Rightarrow b' = \frac{fE(N_{11})}{pE(N_{00})} \tag{8}$$

According to Eqs. (7)(8), g is expressed as follows.

$$g = u - b' = -\frac{f}{p}\ln[\frac{E(N_{00})}{f}] - \frac{fE(N_{11})}{pE(N_{00})} \tag{9}$$

By substituting N_{00} for $E(N_{00})$ and N_{11} for $E(N_{11})$ in Eq. (9), we get the estimator of g as follows.

$$\hat{g} = -\frac{f}{p}\ln(\frac{N_{00}}{f}) - \frac{fN_{11}}{pN_{00}} \tag{10}$$

That is, Eq. (10) exactly specifies how to use the observed N_{00} and N_{11} to estimate the cardinality g of genuine tags. Theorem 1 presents the expectation and the variance of the estimate $\hat{g}$, which are very important to investigate the probabilistic accuracy of the estimator.

Theorem 1. *$\hat{g}$ in Eq. (10) is an unbiased estimator of g, that is, $E(\hat{g}) = g$. The variance of the estimator is $Var(\hat{g}) = \frac{1}{fp^2}e^{\frac{up}{f}}(b'^2p^2 + f^2 - b'fp) - \frac{f}{p^2}$.*

Proof: To get the expectation and variance of $\hat{g}$, we use a similar method in [10]. Since the variance expression is different from [10] [24], we present the detailed proving procedures for the completeness of this paper. According to Eq. (10), $\hat{g}$ is a function with respect to N_{00} and N_{11}. Hence, we denote $\hat{g}$ as $\varphi(N_{00}, N_{11})$, that is, $\hat{g} = \varphi(N_{00}, N_{11})$. We present the Taylor's series expansion [25] of function $\varphi(N_{00}, N_{11})$ around (η_0, η_1), where $\eta_0 = E(N_{00})$ and $\eta_1 = E(N_{11})$.

$$\varphi(N_{00}, N_{11}) \approx \varphi(\eta_0, \eta_1) + [(N_{00} - \eta_0)\frac{\partial\varphi}{\partial N_{00}} + (N_{11} - \eta_1)\frac{\partial\varphi}{\partial N_{11}}] \tag{11}$$

We have the following equation by taking expectation of both sides of Eq. (11).

$$\begin{aligned} &E[\varphi(N_{00}, N_{11})] \\ =&\varphi(\eta_0, \eta_1) + \frac{\partial\varphi}{\partial N_{00}}E(N_{00} - \eta_0) + \frac{\partial\varphi}{\partial N_{11}}E(N_{11} - \eta_1) = g \end{aligned} \tag{12}$$

So far, $E(\hat{g}) = g$ is proved, that is, $\hat{g}$ is an unbiased estimator of g. In what follows, we investigate the variance of $\hat{g}$.

$$\begin{aligned} &Var(\hat{g}) = E[\hat{g} - E(\hat{g})]^2 \\ =&E[(N_{00} - \eta_0)\frac{\partial\varphi}{\partial N_{00}} + (N_{11} - \eta_1)\frac{\partial\varphi}{\partial N_{11}}]^2 \\ =&Var(N_{00})(\frac{\partial\varphi}{\partial N_{00}})^2 + Var(N_{11})(\frac{\partial\varphi}{\partial N_{11}})^2 + \\ &2Cov(N_{00}, N_{11})\frac{\partial\varphi}{\partial N_{00}}\frac{\partial\varphi}{\partial N_{11}} \end{aligned} \tag{13}$$

In the following, we present how to get the covariance $Cov(N_{00}, N_{11})$. Since $Cov(N_{00}, N_{11}) = E(N_{00}N_{11}) - E(N_{00})E(N_{11})$, we calculate $E(N_{00}N_{11})$ below.

$$\begin{aligned} &E(N_{00}N_{11}) = \sum_{x=0}^{f}\sum_{y=0}^{f-x} xyP[N_{00} = x \wedge N_{11} = y] \\ =&\sum_{x=0}^{f}\sum_{y=0}^{f-x} xy\binom{f}{x}(p_{00})^x\binom{f-x}{y}(p_{11})^y(1 - p_{00} - p_{11})^{(f-x-y)} \\ =&p_{11}\sum_{x=1}^{f} f(f-x)\binom{f-1}{x-1}(p_{00})^x(1-p_{00})^{f-x-1} \\ =&\frac{p_{00}p_{11}f^2}{1-p_{00}}\sum_{x=1}^{f}\binom{f-1}{x-1}(p_{00})^{x-1}(1-p_{00})^{f-x} \\ &-\frac{f(f-1)(p_{00})^2p_{11}}{1-p_{00}}\sum_{x=2}^{f}\binom{f-2}{x-2}(p_{00})^{x-2}(1-p_{00})^{f-x} \\ &-\frac{fp_{00}p_{11}}{1-p_{00}}\sum_{x=1}^{f}\binom{f-1}{x-1}(p_{00})^{x-1}(1-p_{00})^{f-x} \\ =&\frac{p_{00}p_{11}f^2}{1-p_{00}} - \frac{f(f-1)(p_{00})^2p_{11}}{1-p_{00}} - \frac{fp_{00}p_{11}}{1-p_{00}} = f(f-1)p_{00}p_{11} \end{aligned} \tag{14}$$

As required by Eq. (13), we also calculate the first-order partial derivatives of $\varphi(N_{00}, N_{11})$ as follows.

$$\begin{aligned} &\frac{\partial\varphi}{\partial N_{00}}\Big|_{N_{11}=\eta_1}^{N_{00}=\eta_0} = e^{\frac{up}{f}}(\frac{b'}{f} - \frac{1}{p}) \\ &\frac{\partial\varphi}{\partial N_{11}}\Big|_{N_{11}=\eta_1}^{N_{00}=\eta_0} = -\frac{1}{p}e^{\frac{up}{f}} \end{aligned} \tag{15}$$

We have obtained $E(N_{00}N_{11})$ in Eq. (14), $E(N_{00})$ in Eq. (2), and $E(N_{11})$ in Eq. (5). Thus, we can calculate $Cov(N_{00}, N_{11})$ as follows.

$$\begin{aligned} &Cov(N_{00}, N_{11}) = E(N_{00}N_{11}) - E(N_{00})E(N_{11}) \\ =&-fp_{00}p_{11} = -b'pe^{-\frac{2up}{f}} \end{aligned} \tag{16}$$

By combining Eqs. (3) (6) (15) (16) into Eq. (13), we then get the variance of $\hat{g}$ as follows.

$$Var(\hat{g}) = \frac{1}{fp^2}e^{\frac{up}{f}}(b'^2p^2 + f^2 - b'fp) - \frac{f}{p^2}, \tag{17}$$

where f and p are the used frame size and the persistence probability, respectively. ■

D. Refined Estimation with k Frames

Because of probabilistic variance, the estimate $\hat{g}$ got from a single frame is hard to meet the predefined accuracy. By the law of large number [26], we issue k independent frames and use the average estimation result $\hat{g_k} = \frac{1}{k}\sum_{j=1}^{k}\hat{g}_j$ to achieve a more accurate estimate in REB, where $\hat{g}_j$ is the estimate of g derived from the j^{th} frame. We propose Theorem 2 to investigate how many independent frames are necessary to guarantee that the average estimate $\hat{g_k}$ can satisfy the predefined (α, β) accuracy.

Theorem 2. *The reader performs k independent frames. The average estimate $\hat{g_k} = \frac{1}{k}\sum_{j=1}^{k}\hat{g}_j$ can guarantee the required (α, β) accuracy, if the frame number k satisfies $k \geq \frac{Z_\beta}{g\alpha}\sqrt{\sum_{j=1}^{k}[\frac{1}{f_jp_j^2}e^{\frac{up_j}{f_j}}(b'^2p_j^2 + f_j^2 - b'f_jp_j) - \frac{f_j}{p_j^2}]}$, where f_j and p_j are the frame size and persistence probability of the j^{th} frame, respectively.*

Proof: We define $\hat{g_k} = \frac{1}{k}\sum_{j=1}^{k}\hat{g}_j$ as the *average* estimate of k successive frames, where $\hat{g}_j$ is the estimate of the j^{th} frame, $j \in [1, k]$. The reader initializes each frame with different random seeds. Hence, the estimate $\hat{g}_j$ is independent to each other. Thus, we have $E(\hat{g_k}) = \frac{1}{k}\sum_{j=1}^{k}E(\hat{g}_j) = g$; and $Var(\hat{g_k}) = \frac{1}{k^2}\sum_{j=1}^{k}Var(\hat{g}_j)$. Clearly, the average estimate $\hat{g_k}$ still converges to the actual cardinality g. Given a required reliability β, the actual confidence interval is within $[g - Z_\beta\sqrt{Var(\hat{g_k})}, g + Z_\beta\sqrt{Var(\hat{g_k})}]$, where Z_β is a percentile of β, *e.g.*, if $\beta = 95\%$, Z_β will be 1.96. To guarantee the required confidence α, we should guarantee:

$$\begin{cases} g + Z_\beta\sqrt{Var(\hat{g_k})} \leq g + g\alpha \\ g - Z_\beta\sqrt{Var(\hat{g_k})} \geq g - g\alpha \end{cases}$$

Substituting $\frac{1}{k^2}\sum_{j=1}^{k}Var(\hat{g}_j)$ for $Var(\hat{g_k})$ and solving the above inequalities, we have:

$$k \geq \frac{Z_\beta}{g\alpha}\sqrt{\sum_{j=1}^{k}Var(\hat{g}_j)} \tag{18}$$

According to Eq. (17), we have $Var(\hat{g}_j) = \frac{1}{f_j p_j^2} e^{\frac{up_j}{f_j}} (b'^2 p_j^2 + f_j^2 - b' f_j p_j) - \frac{f_j}{p_j^2}$. Substituting it into Eq. (18), we have:

$$k \geq \frac{Z_\beta}{g\alpha} \sqrt{\sum_{j=1}^{k} [\frac{1}{f_j p_j^2} e^{\frac{up_j}{f_j}} (b'^2 p_j^2 + f_j^2 - b' f_j p_j) - \frac{f_j}{p_j^2}]} \quad (19)$$

When the above inequality holds, the predefined (α, β) accuracy can be satisfied. ■

After k frames, the backend server calculates the R.H.S. (Right Hand Side) of Eq. (19). If the result is less than (or equal to) k, the estimation process terminates; otherwise, the next frame continues to be issued. Note that, we do not know the actual u, b' and g. We propose to use the first k leading frames to estimate them, as shown in Eq. (20).

$$\begin{aligned} \hat{u}_{\overline{k}} &= \frac{1}{k} \sum_{j=1}^{k} \hat{u}_j = \frac{1}{k} \sum_{j=1}^{k} [-\frac{f_j}{p_j} \ln(\frac{N_{00j}}{f_j})] \\ \hat{b'}_{\overline{k}} &= \frac{1}{k} \sum_{j=1}^{k} \hat{b'}_j = \frac{1}{k} \sum_{j=1}^{k} [\frac{f_j N_{11j}}{p_j N_{00j}}] \\ \hat{g}_{\overline{k}} &= \frac{1}{k} \sum_{j=1}^{k} \hat{g}_j = \frac{1}{k} \sum_{j=1}^{k} [-\frac{f_j}{p_j} \ln(\frac{N_{00j}}{f_j}) - \frac{f_j N_{11j}}{p_j N_{00j}}], \end{aligned} \quad (20)$$

where N_{00j} and N_{11j} are the numbers of persistent empty slots and persistent singleton slots observed in the j^{th} frame.

E. Avoiding Premature Termination

In the execution of REB, we can get $\hat{u}_{\overline{k}}$, $\hat{b'}_{\overline{k}}$ and $\hat{g}_{\overline{k}}$ after k frames. However, their estimation is inaccurate due to probability variance. If we directly use them to calculate the R.H.S. of Eq. (19), k may have a chance to be larger than it, which is not true and REB will have a premature termination (*i.e.*, the currently achieved accuracy has not met the required one yet). In the following, we propose to solve the issue of premature termination. First, we calculate the variances of $\hat{u}_{\overline{k}}$, $\hat{b'}_{\overline{k}}$ and $\hat{g}_{\overline{k}}$ as follows. Note that, we can obtain the variances of $\hat{u}_{\overline{k}}$ and $\hat{b'}_{\overline{k}}$ using similar method of getting $Var(\hat{g}_{\overline{k}})$.

$$\begin{aligned} Var(\hat{u}_{\overline{k}}) &= \frac{1}{k^2} \sum_{j=1}^{k} \frac{f_j}{p_j^2} (e^{\frac{\hat{u}_{\overline{k}} p_j}{f_j}} - 1) \\ Var(\hat{b'}_{\overline{k}}) &= \frac{1}{k^2} \sum_{j=1}^{k} e^{\frac{\hat{u}_{\overline{k}} p_j}{f_j}} (\frac{\hat{b'}^2_{\overline{k}}}{f_j} + \frac{\hat{b'}_{\overline{k}}}{p_j}) \\ Var(\hat{g}_{\overline{k}}) &= \frac{1}{k^2} \sum_{j=1}^{k} \frac{1}{f_j p_j^2} e^{\frac{\hat{u}_{\overline{k}} p_j}{f_j}} (\hat{b'}^2_{\overline{k}} p_j^2 + f_j^2 - \hat{b'}_{\overline{k}} f_j p_j) - \frac{f_j}{p_j^2} \end{aligned} \quad (21)$$

When calculating the R.H.S. of Eq. (19), we can use $\hat{u}_{k\uparrow} = \hat{u}_{\overline{k}} + \delta\sqrt{Var(\hat{u}_{\overline{k}})}$ to substitute u, $\hat{b'}_{k\uparrow} = \hat{b'}_{\overline{k}} + \delta\sqrt{Var(\hat{b'}_{\overline{k}})}$ to substitute the *first* b', $\hat{b'}_{k\downarrow} = \hat{b'}_{\overline{k}} - \delta\sqrt{Var(\hat{b'}_{\overline{k}})}$ to substitute the *second* b', $\hat{g}_{k\uparrow} = \hat{g}_{\overline{k}} + \delta\sqrt{Var(\hat{g}_{\overline{k}})}$ to substitute g. In Section III, simulation results demonstrate that this tactic can effectively avoid the premature termination. The *three-sigma rule* [27] indicates $\delta = 3$ is large enough.

F. Dynamically Optimizing p and f

In REB, parameters p and f can significantly affect the protocol performance, thus need to be optimized. We use the information observed from the x leading frames to facilitate the optimization of p and f in the $(x+1)^{th}$ frame. The optimization goal is to minimize the execution time while guaranteeing the required (α, β) accuracy.

In the first frame, we set $f_1 = 512$. To *coarsely* set p_1, we modify the scheme used in [10], [18], [19]. Specifically, the reader keeps issuing one-slot frames. The persistence probability follows a geometric distribution, $\frac{1}{2}$, $\frac{1}{4}$, $\frac{1}{8}$, $\cdots$, *i.e.*, the persistence probability in the γ^{th} single-slot frame is $\frac{1}{2^\gamma}$. This process does not terminate until an empty slot appears. Assuming the ℓ^{th} slot is the first empty slot, we have a coarse estimation of u to be 2^ℓ [18]. The persistence probability p_1 of the first frame is simply set to $f/2^\ell$.

In what follows, we describe how to optimize p and f for the $(x+1)^{th}$ frame ($x \geq 1$). Since p and f are correlated to minimize the total execution time, we first fix the f value to get an optimized p. The range of f is from 1 to 512 and its value should be 2, 4, 8, $\cdots$, 512, as suggested in the C1G2 standard. Thus, we can get an optimized f by comparing all possible pairs of p and f (with only 9 possible f values). Note that, the proposed REB is *not sensitive* to the coarse settings of p_1 and f_1, because REB will quickly converge to a near-optimal setting of p and f after a few frames, which will be demonstrated in the simulations.

1) Optimizing p: We optimize p using a binary search method for a given f value. Since the smaller the estimation variance of $\hat{g}$ is, the less frames will be required, that is, the less the execution time ($f\times$ frame number) is. We theoretically investigate how to optimize p to *minimize* the estimation variance $Var(\hat{g})$ in Eq. (17).

We prove that $Var(\hat{g})$ is a *convex* function of $p \in (0, 1]$ in Theorem 3. By virtue of the convex property, we have two claims: (i) There is an optimal $p_{op} \in (0, 1]$ minimizing the variance $Var(\hat{g})$. Taking Fig. 2 (a)(b) for example. (ii) The first order partial derivation $\frac{\partial Var(\hat{g})}{\partial p}$ presented in the following Eq. (22) satisfies $\forall p \in (0, p_{op}), \frac{\partial Var(\hat{g})}{\partial p} \leq 0$ and $\forall p \in (p_{op}, 1], \frac{\partial Var(\hat{g})}{\partial p} \geq 0$. Taking Fig. 2 (c)(d) for example.

$$\frac{\partial Var(\hat{g})}{\partial p} = e^{\frac{up}{f}} (\frac{b'^2 u}{f^2} + \frac{b' + u}{p^2} - \frac{b' u}{fp} - \frac{2f}{p^3}) + \frac{2f}{p^3} \quad (22)$$

Based on the above two claims, we propose a *binary-search* algorithm to get the optimal p_{op} for the j^{th} frame, $j \geq 2$. In step 1 of Algorithm 1, δ specifies the maximum deviation between the outputted p and its actually optimal value. Initially, $p_{high} = 1$ and p_{low} is set to a small enough value $1/\hat{u}_{\overline{x}}$. By the while loop in steps 4∼10, p_{high} and p_{low} progressively approach the optimal p_{op}. When the optimization derivation is less than δ, the average value of p_{high} and p_{low} are returned as the optimal p. The complexity of Algorithm 1 is $\Theta(\lg \frac{1}{\delta})$.

Theorem 3. *$Var(\hat{g})$ in Eq. (17) is a convex function of p.*

Proof: $Var(\hat{g})$ is a convex function of $p \in (0, 1]$, *if and only if* its second order partial derivative satisfies $\forall p \in$

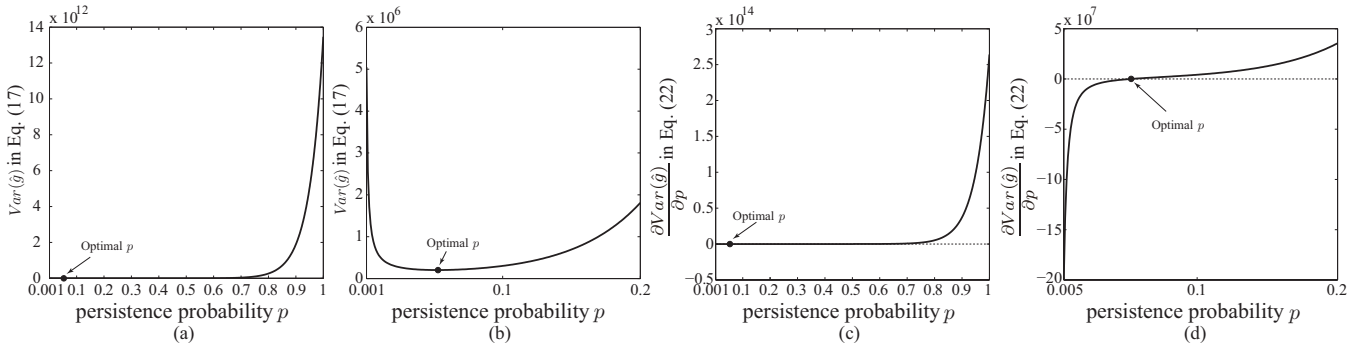

Fig. 2. $b = 10000$, $b' = 5000$, $g = 10000$, $f = 512$. (a) The variance $Var(\hat{g})$ against $p \in [0.001, 1]$. (b) Magnified plot of (a), $p \in [0.001, 0.2]$. (c) The first order partial derivation of $Var(\hat{g})$ against $p \in [0.001, 1]$. (d) Magnified plot of (c), $p \in [0.005, 0.2]$.

$(0, 1]$, $\frac{\partial^2 \hat{g}}{\partial p^2} > 0$. We get its second order partial derivative as follows.

$$\frac{\partial^2 \hat{g}}{\partial p^2} = e^{\frac{up}{f}} \underbrace{(\frac{b'^2u^2}{f^3} + \frac{2b'u + u^2}{fp^2} - \frac{b'u^2}{f^2p} - \frac{2b' + 4u}{p^3} + \frac{6f}{p^4})}_{\text{denoted as } \Re} - \frac{6f}{p^4} \tag{23}$$

As shown in Eq. (23), we denote $(\frac{b'^2u^2}{f^3} + \frac{2b'u+u^2}{fp^2} - \frac{b'u^2}{f^2p} - \frac{2b'+4u}{p^3} + \frac{6f}{p^4})$ as $\Re$. In what follows, we first prove $\Re$ is always larger than 0.

$$\begin{aligned}\Re =& (b'^2u^2p^4 + 2b'uf^2p^2 + u^2f^2p^2 - b'u^2fp^3 - 2b'f^3p \\ & - 4uf^3p + 6f^4)/(f^3p^4) \\ =& [(b'up^2 - \frac{1}{2}ufp)^2 + (\sqrt{2b'u}fp - \frac{\sqrt{2}}{2}f^2)^2 + \frac{1}{6}f^4 \\ & + (\frac{\sqrt{3}}{2}ufp - \frac{4\sqrt{3}}{3}f^2)^2 + 2(\sqrt{b'u} - b')f^3p]/(f^3p^4)\end{aligned} \tag{24}$$

Since $u > b'$ in Eq. (24), we have $\Re > 0$. Using the fourth-order Taylor series expansion, we have $e^{\frac{up}{f}} > 1 + \frac{up}{f} + \frac{u^2p^2}{2f^2} + \frac{u^3p^3}{6f^3} + \frac{u^4p^4}{24f^4}$. According to Eqs. (23)(24), we have:

$$\begin{aligned}\frac{\partial^2 \hat{g}}{\partial p^2} =& e^{\frac{up}{f}}\Re - \frac{6f}{p^4} \\ >& (1 + \frac{up}{f} + \frac{u^2p^2}{2f^2} + \frac{u^3p^3}{6f^3} + \frac{u^4p^4}{24f^4})\Re - \frac{6f}{p^4} \\ =& \frac{1}{24}(\frac{u^3p}{2f^2\sqrt{f}} - \frac{u^3b'p^2}{f^3\sqrt{f}})^2 + \frac{5u^6p^2}{1152f^5} + \frac{p^2}{2f^5}(u^2b' - \frac{u^3}{12})^2 \\ & + \frac{1}{f^3}(\sqrt{\frac{2}{3}}ub' - \sqrt{\frac{3}{128}}\frac{u^3p}{f})^2 + \frac{u^2}{3f^3}(b' - \frac{u}{2})^2 \\ & + \frac{u^3b'^2p}{f^4} + \frac{u^5b'^2p^3}{6f^6} + \frac{2u - 2b'}{p^3} > 0\end{aligned} \tag{25}$$

Eq. (25) indicates that $\forall p \in (0, 1]$, $\frac{\partial^2 \hat{g}}{\partial p^2} > 0$, which is a necessary and sufficient condition to prove that $Var(\hat{g})$ is a *convex function* of $p \in (0, 1]$. ■

2) Optimizing f: We optimize f for the $(x+1)^{th}$ frame. Considering C1G2 standard and practical constraints [10], f should take a value from 2, 4, 8, $\cdots$, or 512. To improve the time-efficiency, it is reasonable to minimize the expected remaining execution time. We denote the minimum frame number that needs to be further executed as y, our goal is:

$$\text{Minimizing } (f + 1) \times y \tag{26}$$

Algorithm 1: Optimizing p_{x+1} for the $(x+1)^{th}$ frame.

Input: $\hat{u}_{\overline{x}}$, $\hat{b'}_{\overline{x}}$, $\hat{g}_{\overline{x}}$, and f.
Output: The optimized p_{x+1} for the $(x+1)^{th}$ frame.

```
δ = 0.0001;
p_low = 1/û_x̄;
p_high = 1;
while p_high − p_low > δ do
    p = (p_low + p_high)/2;
    Calculating ∂Var(ĝ)/∂p in Eq. (22);
    if (∂Var(ĝ)/∂p > 0) then
        p_high = p;
    else
        p_low = p;
    end if
end while
p_{x+1} = (p_low + p_high)/2;
return p_{x+1};
```

In Eq. (26), $f + 1$ means a slot for transmitting protocol parameters is followed by an f-slot frame. According to the termination condition in Eq. (18), the value of y should satisfy the following inequality:

$$x + y \geq \frac{Z_\beta}{g\alpha}\sqrt{\sum_{j=1}^{x} Var(\hat{g}_j) + yVar(\hat{g})} \tag{27}$$

By solving the above inequality, we know that the minimum frame number y that needs to be further executed is:

$$\lceil \frac{Z_\beta^2 Var(\hat{g})}{2g^2\alpha^2} - x + \sqrt{[\frac{Z_\beta^2 Var(\hat{g})}{2g^2\alpha^2} - x]^2 + \frac{Z_\beta^2}{g^2\alpha^2}\sum_{j=1}^{x} Var(\hat{g}_j) - x^2} \rceil \tag{28}$$

Here, $Var(\hat{g}_j) = \frac{1}{f_j p_j^2} e^{\frac{up_j}{f_j}}(b'^2p_j^2 + f_j^2 - b'f_jp_j) - \frac{f_j}{p_j^2}$, $j \in [1, x]$, $Var(\hat{g}) = \frac{1}{fp^2}e^{\frac{up}{f}}(b'^2p^2 + f^2 - b'fp) - \frac{f}{p^2}$. g, u, b' can be approximated by $\hat{g}_{\overline{x}}$, $\hat{u}_{\overline{x}}$, $\hat{b'}_{\overline{x}}$, respectively.

For any $f \in \{2, 4, 8, \cdots, 512\}$, we can use Algorithm 1 to get the corresponding optimal p. Given each pair (f, p), the smallest y can be obtained from Eq. (28). Thus, we can get the optimal f from Eq. (26). The calculation complexity of optimizing p and f is bounded by $\Theta(9 \lg \frac{1}{\delta})$.

G. REB with Multiple Readers and Blocker Tags

In practice, a single reader cannot probe all the tags due to the limited communication ranges [10]. Similarly, a single blocker tag cannot "protect" all privacy-sensitive tags that may be distributed across the a large area. A solution is to deploy multiple readers and blocker tags with overlapping regions to cover the whole monitoring area. We assume all the readers and blocker tags are well synchronized by the excellent scheduling schemes [28]–[30]. All parameters f, p, R involved in REB are the same across all readers. In what follows, we present how to distributively construct the *global* $\mathbb{BG}[0..f-1]$. For an arbitrary slot s in a frame, if all the readers observe an empty slot, the backend server sets $\mathbb{BG}[s] = 0$; if no reader senses a collision *and* all the received RN16s are the same, the backend server sets $\mathbb{BG}[s] = 1$; if at least one reader senses a collision *or* different RN16s are observed by different readers, the backend server sets $\mathbb{BG}[s] = c$. Based on these rules, the reader is able to generate a global actual array $\mathbb{BG}[0..f-1]$. Logically, all the readers co-work like a 'super reader' that is able to cover the whole area. The rest of REB in multi-reader scenarios is the same as what former sections have described.

III. Performance Evaluation

In this section, we conduct simulations to evaluate the performance of REB in a large scale RFID system that contains thousands of tags. The simulators were implemented using MATLAB. Since REB in the multi-reader scenario is logically the same as that in the single reader scenario, this paper only simulates a single reader following prior literature [7], [10], [16], [31]. The setting of slot length is based on what we specified in Section II-A. In the following, we first verify the effectiveness of our optimization methods on f and p. Then, we conduct simulations to evaluate the actual estimation reliability of REB and its time-efficiency. We run each simulation 1000 times and report the average results.

A. Verifying the Optimized f and p

The setting of parameters f and p is important to the performance of REB. To achieve the *overall* optimal f and p, it is necessary to know the values of u, b' and g before the execution of REB, which, however, are what we want to estimate. Using the simulation conditions shown in Fig. 3, the *overall* optimal f is 128, and the *overall* optimal p is 0.01175, which are calculated by the actual values of u, b' and g. In the simulations corresponding to Fig. 3, we aim to verify the convergence of f and p to their *overall* optimal values. Results in Fig. 3 (a) demonstrate that about 28.5% of the independent simulations correctly take the *overall* optimal $f = 128$ in the 3^{rd} frame. And this ratio reaches 50.5% in the 4^{th} frame, that is, our REB has a good chance to take the overall optimal f just after the 3^{rd} frame. The simulation results in Fig. 3 (b) demonstrate that the persistent probability p approaches its overall optimal value frame by frame. The value of p taken in the 4^{th} is very close to the optimal value 0.01175. All in all, f and p approach their *overall* optimal values frame by frame. The underlying reason is that more frames increase the estimation accuracy of u, b' and g, which eventually facilitates the optimization of f and p.

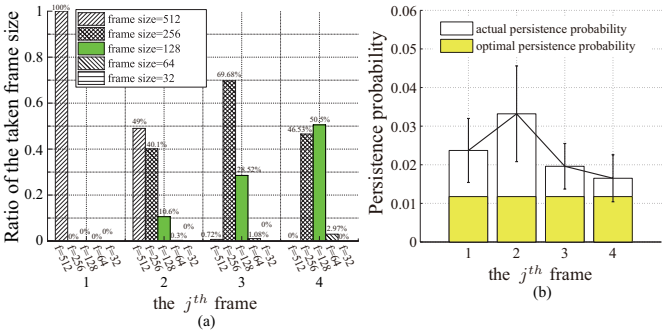

Fig. 3. Verifying the optimized settings of f and p. $|B-G| = 5000$, $|B \cap G| = 5000$, $|G-B| = 5000$. $\alpha = 10\%$, $\beta = 90\%$. (a) Verifying the optimized f. (b) Verifying the optimized p.

B. Estimation Reliability

One of the most important performance metrics for estimation protocols is the actual reliability. In an arbitrary simulation, if the estimate $\hat{g}$ is within $[g(1-\alpha), g(1+\alpha)]$, we refer to it as a successful estimation. We record the *success times* among 1000 independent simulations. The ratio, *i.e.*, *success times*/1000, is treated as the *actual reliability*. Simulation results in Fig. 4 reveal that REB ($\delta = 0$) does not always meet the required reliability (*i.e.*, $\beta = 95\%$). The reason lies in the variances if directly using $\hat{u}_{\overline{k}}$, $\hat{b'}_{\overline{k}}$ and $\hat{g}_{\overline{k}}$ to determine the termination condition. By taking their variances into consideration, the proposed δ-sigma-based termination tactic effectively avoids the premature termination. Simulation results in Fig. 4 reveal that the actual reliability of REB ($\delta = 1$) and REB ($\delta = 2$) is always higher than the required one.

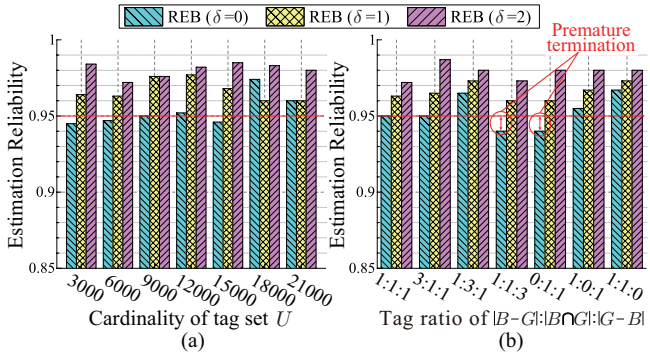

Fig. 4. Evaluating the reliability of REB. $\alpha = 5\%$, $\beta = 95\%$. (a) Tag ratio $|B-G|:|B \cap G|:|G-B|$ is fixed to $1:1:1$, and u varies from 3000 to 21000. (b) u is fixed to 9000, and tag ratio varies.

C. Time Efficiency

Besides the estimation reliability, another important metric is time-efficiency. In this subsection, we evaluate the time-efficiency of the protocols *given the same estimation accuracy*. No existing estimation protocols can correctly approximate the cardinality of genuine tags in an RFID system with the presence of blocker tags. The only possible solution, to the best of our knowledge, is to perform the comprehensive identification protocols to identify the tags in the system. Hence, we compare REB with two representative identification protocols: one is the Tree Hopping

(TH) protocol [19]; the other one is the Enhanced Dynamic Framed Slotted ALOHA (EDFSA) protocol [20]. TH protocol terminates after it traverses the whole tree. TH can identify not only the IDs in $(B-G)\cup(G-B)$ when a queried prefix is followed by a successful read, but also the IDs in $B\cap G$ when a prefix whose length is equal to tag ID but still followed by a collision read. Then, we can get the set G, by calculating $[(B-G)\cup(G-B)-B]\cup(B\cap G)$. The cardinality g is got upon getting G. As for EDFSA protocol, it executes frames round by round. In a round, only the IDs in $(B-G)\cup(G-B)$ have chance to be identified. We denote the set of identified IDs as S_{ident}. Since the reader does not know whether all IDs in $(B-G)\cup(G-B)$ are completely identified or even what percentage of them are identified, EDFSA cannot terminate by itself. For the sake of EDFSA, we assumes it can "intelligently" terminate once $[|(B-G)\cup(G-B)|-|S_{ident}|]<|G|\times\alpha$.

1) Impact of Tag Cardinality: To investigate the impact of tag cardinality on the protocols' execution time, we fix the tag ratio $|B-G|:|B\cap G|:|G-B|$ to 1:1:1, and vary u (indicating the system scale) from 20000 to 50000. The simulation results in Fig. 5 demonstrate that our REB significantly outperforms HT and EDFSA. For example, when $u=50000$, REB ($\delta=0$) runs about 44 times faster than EDFSA, and nearly 920 times faster than TH; while REB ($\delta=1$) runs 33 times faster than EDFSA, and 682 times faster than TH. Moreover, the execution time of HT and EDFSA grows linearly as u increases. In contrast, our REB has a stable execution time, which reveals its good scalability against tag cardinality u.

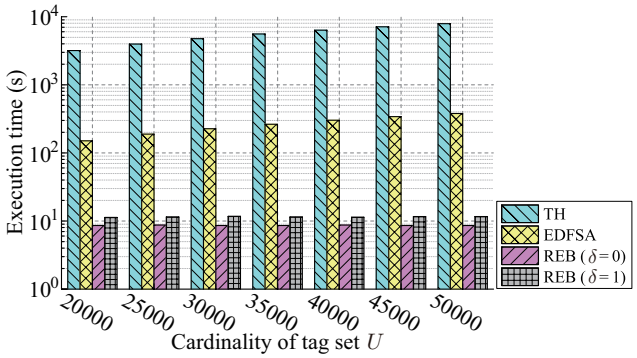

Fig. 5. Evaluating the time-efficiency of protocols with varying u. Tag ratio of $|B-G|:|B\cap G|:|G-B|$ is fixed to $1:1:1$ and $\alpha=5\%$, $\beta=95\%$.

2) Impact of Tag Ratio: The different tag ratio of $|B-G|:|B\cap G|:|G-B|$ may have significant impact on the execution time of protocols. Here, we fix $u=30000$, and evaluate the execution time of protocols with varying tag ratio. The simulation results in Fig. 6 demonstrate that our REB still outperforms the existing protocols by significantly reducing the execution time. Moreover, the results in Fig. 6 clearly show the performance trend of the protocols with varying tag ratio, which are elaborated below.

The results in Fig. 6 (a) reveal that the larger the ratio of tags in $B-G$ is, the larger the execution time of our REB scheme is. The underlying reason is that more tags in the set $B-G$ will incur more interferences to the process of estimating genuine tags. We make another two main observations from Fig. 6 (b) which shows the execution time of the protocols with varying ratio of tags in $B\cap G$. First, the performance of identification protocols deteriorates as the ratio of tags in $B\cap G$ increases. The reason is that more tags in $B\cap G$ will cause more blocking collisions, which seriously interfere the tag identification process. Second, the larger the ratio of tags in $B\cap G$ is, the smaller the execution time of our REB scheme is. The underlying intuitive reason is that larger ratio of tags in $B\cap G$ leads to smaller ratio of tags in $B-G$, which decreases the interference of tags in $B-G$ to the process of estimating g. Because of a similar reason, the results in Fig. 6 (c) reveal that the execution time of REB also decreases as the ratio of tags in $G-B$ increases.

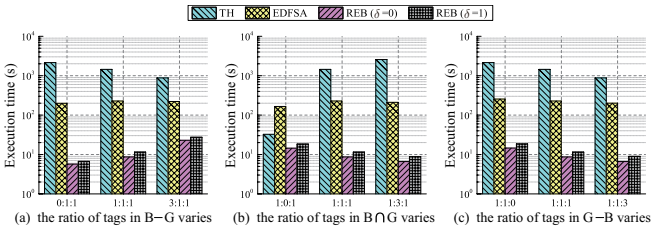

Fig. 6. Evaluating the time-efficiency of protocols with varying tag ratio. u is fixed to 30000, and $\alpha=5\%$, $\beta=95\%$.

IV. Related Work

In the infant stage of RFID study, a great deal of attention was paid to the problem of tag identification that aims to identify the exact tag IDs. Generally, there are two types of identification protocols: Aloha-based protocols [32] and Tree-based protocols [19]. Fundamentally, the Aloha-based protocol is a kind of Time Division Multiple Access (TDMA) mechanism. A tag ID can be successfully identified in a slot when only one tag responds in this slot. As for tree-based protocols, the reader broadcasts a 0/1 string to query the tags. A tag responds with its ID once it finds that the querying string is the prefix of its ID. A reader can successfully identify a tag ID when only one tag responds. Clearly, the execution time of identification protocols is proportional to the tag population size. What is worse, in the RFID system with presence of blocker tags, the performance of identification protocols will further deteriorate because of the blocking collisions caused by IDs in $B\cap G$.

To fast report the tag cardinality for various purposes such as timely stock monitoring, a great effort has been made to study the problem of tag estimation [5], [7], [10], [13]–[18], [23], [33]–[35]. These estimation protocols leverage the observations from Aloha/Tree protocols to statistically estimate the tag cardinality. For example, M. Shahzad *et al.* proposed the Average Run based Tag estimation (ART) by observing the average length of sequences of consecutive non-empty slots [10]. To the best of our knowledge, all these estimation protocols cannot address the problem of genuine tag estimation because they cannot exclude the interference from blocking tags.

RFID privacy is of great importance but suffers the threat from malicious scanning. Ari Juels *et al.* proposed the blocker tags to protect the privacy-sensitive tags from

malicious scanning [11]. Every coin has two sides. Ehsan Vahedi *et al.* indicated that the blocking technique causes a new threat to the RFID system. Specifically, a malicious blocker tag can prevent the valid reader from reading the tags. An efficient scheme was proposed to detect the existence of an attacker in the RFID system [36]. Following the original purpose of proposing blocker tags, this paper still leverages the blocker tags to protect the privacy of genuine RFID tags.

V. Conclusion

This paper formally defines a new problem of genuine tag cardinality estimation with the presence of blocker tags. To efficiently address this practically important problem, we propose the *RFID Estimation scheme with Blocker tags (REB)*, which is compliant with the commodity EPC C1G2 standard and does not require any modifications to off-the-shelf RFID tags. REB provides an unbiased functional estimator which can guarantee any degree of estimation accuracy specified by the users. Using REB, a retailer can timely monitor the product stock while blocker tags are being used to protect the privacy of some important items. Extensive simulation results reveal that REB is tens of times faster than the fastest identification protocol with the same accuracy requirement.

Acknowledgment

This work is supported by the National Science Foundation for Distinguished Young Scholars of China (Grant No. 61225010); the State Key Program of National Natural Science of China(Grant No. 61432002); NSFC under Grant nos. of 61173161, 61173162, 61272417, 61300187, 61300189, 61370198, 61370199, 61472184, 61321491 and 61272546; HK RGC PolyU G-YM08; NSF grants ECCS 1231461, ECCS 1128209, CNS 1138963, CNS 1065444, and CCF 1028167; the Jiangsu Future Internet Program under Grant No. BY2013095-4-08, and the Jiangsu High-level Innovation and Entrepreneurship (Shuangchuang) Program.

References

[1] K. Finkenzeller, "RFID Handbook: Fundamentals and Applications in Contactless Smart Cards, Radio Frequency Identification and Near-Field Communication," *Wiley*, 2010.

[2] L. Yang, Y. Chen, X.-Y. Li, C. Xiao, M. Li, and Y. Liu, "Tagoram: Real-Time Tracking of Mobile RFID Tags to High Precision Using COTS Devices," *Proc. of ACM MobiCom*, 2014.

[3] S. Qi, Y. Zheng, M. Li, L. Lu, and Y. Liu, "COLLECTOR: A Secure RFID-Enabled Batch Recall Protocol," *Proc. of IEEE INFOCOM*, 2014.

[4] T. Liu, L. Yang, Q. Lin, and Y. Liu, "Anchor-free Backscatter Positioning for RFID Tags with High Accuracy," *Proc. of IEEE INFOCOM*, 2014.

[5] L. Xie, H. Han, Q. Li, J. Wu, and S. Lu, "Efficiently Collecting Histograms Over RFID Tags," *Proc. of IEEE INFOCOM*, 2014.

[6] J. Liu, B. Xiao, K. Bu, and L. Chen, "Efficient Distributed Query Processing in Large RFID-enabled Supply Chains," *Proc. of IEEE INFOCOM*, 2014.

[7] Y. Zheng and M. Li, "PET: Probabilistic Estimating Tree for Large-scale RFID Estimation," *IEEE Transactions on Mobile Computing*, vol. 11, no. 11, pp. 1763–1774, 2012.

[8] M. Roberti, "A 5-cent Breakthrough," *RFID Journal*, vol. 5, no. 6, 2006.

[9] "http://www.centreforaviation.com/news/share-market/2010/06/17/hong-kong-airport-sets-new-cargo-traffic-record-fedex-sees-surging-asian-exports/page1."

[10] M. Shahzad and A. X. Liu, "Every Bit Counts: Fast and Scalable RFID Estimation," *Proc. of ACM MobiCom*, 2012.

[11] A. Juels, R. L. Rivest, and M. Szydlo, "The Blocker Tag: Selective Blocking of RFID Tags for Consumer Privacy," *Proc. of ACM CCS*, 2003.

[12] "http://www.informationweek.com/rsa-unveils-rfid-tag-blocker/d/d-id/1023433?"

[13] M. Kodialam and T. Nandagopal, "Fast and Reliable Estimation Schemes in RFID Systems," *Proc. of ACM Mobicom*, 2006.

[14] M. Kodialam, T. Nandagopal, and W. C. Lau, "Anonymous Tracking using RFID tags," *Proc. of IEEE INFOCOM*, 2007.

[15] C. Qian, H. Ngan, Y. Liu, and L. M. Ni, "Cardinality Estimation for Large-scale RFID Systems," *IEEE Transactions on Parallel and Distributed Systems*, vol. 22, no. 9, pp. 1441–1454, 2011.

[16] B. Chen, Z. Zhou, and H. Yu, "Understanding RFID Counting Protocols," *Proc. of ACM MobiCom*, 2013.

[17] Y. Zheng and M. Li, "ZOE: Fast Cardinality Estimation for Large-Scale RFID Systems," *Proc. of IEEE INFOCOM*, 2013.

[18] T. Li, S. Wu, S. Chen, and M. Yang, "Energy Efficient Algorithms for the RFID Estimation Problem," *Proc. of IEEE INFOCOM*, 2010.

[19] M. Shahzad and A. X. Liu, "Probabilistic Optimal Tree Hopping for RFID Identification," *Proc. of ACM SIGMETRICS*, 2013.

[20] S. Lee, S. Joo, and C. Lee, "An Enhanced Dynamic Framed Slotted ALOHA Algorithm for RFID Tag Identification," *Proc. of IEEE MobiQuitous*, 2005.

[21] P. Semiconductors, "I-CODE Smart Label RFID Tags," *http://www.nxp.com/acrobat_download/other/identification/SL092030.pdf*, 2004.

[22] E. Inc, "Radio-frequency Identity Protocols Class-1 Generation-2 UHF RFID Protocol for Communications at 860 mhz-960 mhz," *EPCGlobal*, Inc, 1.2.0 ed., 2008.

[23] W. Gong, K. Liu, X. Miao, Q. Ma, Z. Yang, and Y. Liu, "Informative Counting: Fine-grained Batch Authentication for Large-Scale RFID Systems," *Proc. of ACM MobiHoc*, 2013.

[24] X. Liu, K. Li, H. Qi, B. Xiao, and X. Xie, "Fast Counting the Key Tags in Anonymous RFID Systems," *Proc. of IEEE ICNP*, 2014.

[25] D. E. Smith, "A source book in mathematics," *Courier Dover Publications*, 2012.

[26] M. Schilling, "Understanding Probability: Chance Rules in Everyday Life," *The American Statistician*, vol. 60, no. 1, pp. 97–98, 2006.

[27] S. N. V and D.-B. IV, "Mathematische Statistik in der Technik," *Deutscher Verl. der Wissenschaften*, 1963.

[28] L. Yang, J. Han, Y. Qi, C. Wang, T. Gux, and Y. Liu, "Season: Shelving Interference and Joint Identification in Large-Scale RFID Systems," *Proc. of IEEE INFOCOM*, 2011.

[29] S. Tang, J. Yuan, M. Li, G. Chen, Y. Liu, and J. Zhao, "Raspberry: A Stable Reader Activation Scheduling Protocol in Multi-reader RFID Systems," *Proc. of IEEE ICNP*, 2009.

[30] J. Waldrop, D. W. Engels, and S. E. Sarma, "Colorwave: An Anticollision Algorithm for the Reader Collision Problem," 2003.

[31] T. Li, S. Chen, and Y. Ling, "Identifying the Missing Tags in a Large RFID System," *Proc. of ACM MobiHoc*, 2010.

[32] F. C. Schoute, "Dynamic Frame Length ALOHA," *IEEE Transactions on Communications*, vol. 31, no. 4, pp. 565–568, 1983.

[33] W. Gong, K. Liu, X. Miao, and H. Liu, "Arbitrarily Accurate Approximation Scheme for Large-Scale RFID Cardinality Estimation," *Proc. of IEEE INFOCOM*, 2014.

[34] H. Liu, W. Gong, L. Chen, W. He, K. Liu, and Y. Liu, "Generic Composite Counting in RFID Systems," *Proc. of IEEE ICDCS*, 2014.

[35] Q. Xiao, B. Xiao, and S. Chen, "Differential Estimation in Dynamic RFID Systems," *Proc. of IEEE INFOCOM*, 2013.

[36] E. Vahedi, V. Shah-Mansouri, V. W. S. Wong, I. F.Blake, and R. K. Ward, "Probabilistic Analysis of Blocking Attack in RFID Systems," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 803–817, 2011.